PLoS ONE, 2014, vol.9, N12

Universal internucleotide statistics in full genomes: A footprint of the DNA structure and packaging?

Bogachev M., Kayumov A., Bunde A.
Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© 2014 Bogachev et al. Uncovering the fundamental laws that govern the complex DNA structural organization remains challenging and is largely based upon reconstructions from the primary nucleotide sequences. Here we investigate the distributions of the internucleotide intervals and their persistence properties in complete genomes of various organisms from Archaea and Bacteria to H. Sapiens aiming to reveal the manifestation of the universal DNA architecture. We find that in all considered organisms the internucleotide interval distributions exhibit the same q -exponential form. While in prokaryotes a single q-exponential function makes the best fit, in eukaryotes the PDF contains additionally a second q -exponential, which in the human genome makes a perfect approximation over nearly 10 decades. We suggest that this functional form is a footprint of the heterogeneous DNA structure, where the first q-exponential reflects the universal helical pitch that appears both in pro- and eukaryotic DNA, while the second q -exponential is a specific marker of the large-scale eukaryotic DNA organization.

http://dx.doi.org/10.1371/journal.pone.0112534